MILP Formulations for Generator Maintenance Scheduling in Hydropower Systems

Citation for published version:
Rodriguez, JA, Anjos, MF, Cote, P & Desaulniers, G 2018, 'MILP Formulations for Generator Maintenance Scheduling in Hydropower Systems', *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6171-6180. https://doi.org/10.1109/TPWRS.2018.2833061

Digital Object Identifier (DOI):
10.1109/TPWRS.2018.2833061

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
IEEE Transactions on Power Systems

MILP Formulations for Generator Maintenance Scheduling in Hydropower Systems

Jesus A. Rodriguez, Miguel F. Anjos, *Member,* Pascal Côté, and Guy Desaulniers

***Abstract*—Maintenance activities help prevent costly generator breakdowns but because generators under maintenance are typically unavailable, the impact of maintenance schedules is significant and their cost must be accounted for when planning maintenance. In this paper we address the generator maintenance scheduling problem in hydropower systems. We propose a mixed-integer programming model that considers the time windows of the maintenance activities, as well as the nonlinearities and disjunctions of the hydroelectric production functions. Because the resulting model is hard to solve, we also propose an extended formulation, a set reduction approach that uses logical conditions for excluding unnecessary set elements from the model, and valid inequalities. We performed computational experiments using a variety of instances adapted from a real hydropower system in Canada, and the extended formulation with set reduction achieved the best results in terms of computational time and optimality gap.**

***Index Terms*—Hydroelectric power generation, Integer linear programming, Mathematical programming, Optimal maintenance scheduling**

I. Notation

We denote decision variables and indices with lowercase, and parameters with uppercase.

Primary sets

$\mathcal{I}$: Powerhouses

$\mathcal{M}$: Maintenance tasks

$\mathcal{T}$: Planning time periods, $t \in \mathcal{T} = \{1 \dots T\}$.

Parameters

B_t^-, B_t^+ : Prices of electricity purchase and sale, respectively, at time period t, where $B_t^- \geq B_t^+$ [\$/MWh].

C_{mt} : Total cost of maintenance task m started at time period t [\$].

D_m : Duration of maintenance task m [day].

E_m : Earliest start time period of maintenance task m.

F_{it} : Lateral inflows to powerhouse i in period t [m^3/s].

$\bar{G}_{it}, \underline{G}_i$: Limits on the number of available turbines in powerhouse i at time period t [turbines].

J_t : Electricity load at time t [MWh].

L_m : Latest start time period of maintenance task m.

O_{it} : Maximum number of turbine outages in powerhouse i at time period t [turbines].

$\bar{P}_i$: Generation capacity in powerhouse i [MWh/day].

$\bar{P}_{ik}$: Generation capacity in powerhouse i when k turbines are active [MWh/day].

Q : Factor for conversion from m^3/s to hm^3/day [0.0864·s·hm^3 ·/(day·m^3)].

$\bar{R}_{it}$: Number of maintenance activities that *can* be in execution at powerhouse i in time period t.

$\underline{R}_{it}$: Number of maintenance activities that *must* be in execution at powerhouse i in time period t.

S_{0i} : Initial stored water in reservoir of powerhouse i [hm^3].

$\underline{S}_i, \bar{S}_i$: Limits on stored water in reservoir of powerhouse i [hm^3].

$\bar{U}_i$: Maximum discharge rate in powerhouse i [m^3/s].

$\bar{V}_i$: Maximum water spill in powerhouse i [m^3/s].

$\bar{W}_t^+$: Maximum electricity sale at time t [MWh/day].

$\bar{W}_t^-$: Maximum electricity purchase at time t [MWh/day].

Derived sets

$\mathcal{T}(m)$: Time periods when maintenance task m can be initiated in order to be completed within $\mathcal{T}$.

$\mathcal{M}(i)$: Maintenance tasks m that should be executed in powerhouse i.

$\mathcal{M}(i,t)$: Maintenance tasks m that can be in execution in powerhouse i at time period t.

$\mathcal{U}(i)$: Powerhouses upstream of powerhouse i ($\mathcal{U}(i) \subset \mathcal{I}$).

$\mathcal{K}(i,t)$: Numbers of generators that can be active at time period t and powerhouse i.

$\mathcal{H}(i,k)$: Hyperplanes for approximating the maximum power output of powerhouse i when k turbines are active.

Parameters with indexes in derived sets

β_h^u : Coefficient of u_{it} in hyperplane $h \in \mathcal{H}(i,k)$ for bounding the power output of powerhouse i when k generators are active [MWh· s/(m^3·day)].

β_h^s : Coefficient of s_{it} in hyperplane $h \in \mathcal{H}(i,k)$ for bounding the power output of powerhouse i when k generators are active [MWh/(hm^3·day)].

β_h^0 : Independent term of hyperplane $h \in \mathcal{H}(i,k)$ for bounding the power output of powerhouse i when k generators are active [MWh/day].

Decision variables

r_{it} : Number of maintenance activities in execution at powerhouse i and time period t.

p_{it} : Corrected estimate of the electricity production of powerhouse i during time period t [MWh/day].

$\hat{p}_{it}$: Outer approximation of the electricity production in of powerhouse i during time period t [MWh/day].

p_{itk} : Estimate of the electricity production in powerhouse i during time period t when k generators are active [MWh/day].

s_{it} : Content of reservoir in powerhouse i at the end of period t [hm^3].

v_{it} : Water spill of reservoir in powerhouse i at time period t [m^3/s].

u_{it} : Water discharge of turbines in powerhouse i at time period t [m^3/s].

w_t^-, w_t^+ : Purchase and sale of electricity, respectively, at period t [MWh].

y_{mt} : Binary variable with value 1 if maintenance task m initiates at time period t, 0 otherwise.

z_{itk} : Binary variable with value 1 if k hydro-turbines are active in powerhouse i at time t, 0 otherwise.

II. Introduction

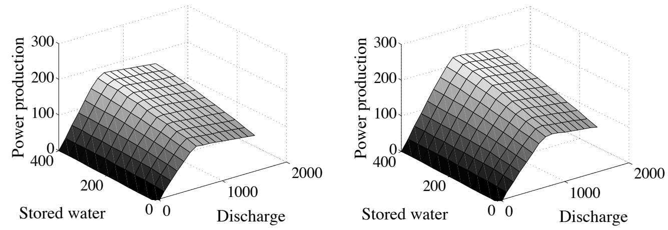
In the power industry, preventive maintenance activities are carried out on a regular basis to prevent expensive equipment failures and to ensure a continuous operation within acceptable efficiency levels. As generators under maintenance are typically inactive, turbine discharges, water spill and electricity production are affected by maintenance activities. Therefore, the maintenance scheduler should anticipate the impact of the maintenance plan on the system operation cost. However in hydroelectric systems these costs are difficult to estimate due to multiple interrelated physical variables. In particular, hydroelectricity production is a function of both the potential energy (the water head) and the kinetic energy of the water that drives the turbine-generators of the system. The relationship between these variables is defined by the Hydropower Production Function (HPF)

$$p = \rho g \gamma q h \eta(q, h), \tag{1}$$

where p is the power output (MW), ρ the water density (kg/m^3), g the gravitational acceleration (m/s^2), γ the conversion factor (10^{-6}), q the turbine water discharge (m^3/s), h the net water head (m), and $\eta(q, h)$ the turbine-generator efficiency (%). For each turbine the efficiency η is a nonlinear function of the net water head and the water discharge of the turbine. Therefore, the efficiency factor η introduces further nonlinearities in the power production of the system.

As the set of active generators affects the generation capacity as well as the optimal quantities of water spill and water discharge, the number of active generators has a nonlinear effect on the total power output. Fig. 1 shows the power production as a function of water discharge and stored water in a reservoir for either four or five active generators.

Spatial and temporal inter-dependences should also be considered in the hydropower operation. First, because water discharges can feed downstream turbines in the current or in subsequent time periods, and second, because future operation costs are determined by present decisions, such as generator outages and water spills from reservoirs. All the aforementioned elements make the optimal planning of maintenance outages in hydropower systems a challenging endeavor.



(a) Four generators (b) Five generators

Fig. 1: The maximum power output as a function of water discharge and stored water varies according to the number of active generators

In the electricity industry, the Generator Maintenance Scheduling Problem (GMSP) has been widely studied, see e.g. [1]. However, specific operating conditions of hydroelectric systems have been scarcely addressed in GMSP. In a GMSP, Feng et al. [2] represented the uncertainty of the power output with fuzzy variables, but omitted water storage levels and water head effects. Foong et al. [3] proposed a meta-heuristic for a GMSP with oversimplified hydropower operation that considers constant power output in active units. Kuzle et al. [4] introduced transmission constraints in a simple GMSP where the nonlinearity of the production functions is neglected. Likewise, Perez-Canto [5] omitted relevant characteristics of hydropower systems, such as temporal and spatial interdependencies, and nonlinearities of the power production. Clearly, a finer representation of the hydropower system's characteristics is necessary to achieve valid solutions to the GMSP for hydropower systems in practice.

Several works have considered the nonlinearity of the HPF (1) for short-term hydropower operation, without incorporating maintenance scheduling decisions [6]-[12]. Finardi and Silva [6], Arce et al. [7] and Catalão et al. [8] used nonlinear functions for estimating the power production of hydro units. However, as the nonlinearity of the HPF makes hydrower scheduling problems hard to solve, different linear approximation approaches have been proposed. For the day-ahead scheduling of generators, Conejo et al. [9] introduced piecewise linearization for representing the effects of the water discharge on the power production. The water head effect on the power output was estimated by interpolation among piecewise approximations for different stored water levels. Following a similar approach, Borghetti et al. [10] proposed a refined linearization for representing the water head effects in hydro unit commitment. Due to the size of the resulting model, results were only reported for a single-reservoir system. For the short-term hydrothermal dispatch problem, Diniz and Piñeiro [11] approximated with linear inequalities the HPF (1), considering the effects of water spill and water head. More recently, Seguin et al. [12]

approximated the power output with smoothing splines for the short-term scheduling of hydro units. These splines were fitted to a maximum power output surface computed by means of dynamic programming for different values of water discharge and stored water level in a reservoir.

In this paper, we propose a mixed-integer linear programming model for the GMSP in hydropower systems that accounts for the nonlinearities of hydroelectric operations via a convex hull approximation of the hydropower production function. Given the difficulty of the resulting optimization problem, we explore three approaches for strengthening the formulation: extended formulation, set reduction, and valid inequalities. The set reduction uses logical conditions for excluding superfluous set elements, in order to reduce the variables and constraints of the model. The possible combinations of these approaches lead to eight formulations that we compared in terms of computational times and optimality gaps on test instances adapted from a real hydropower system in Canada.

This paper is structured as follows. Section II presents our basic mixed integer programming mathematical model. Section III describes the approaches to improve the formulation and the resulting alternative formulations. Section IV reports our computational experiments for evaluating the different alternatives. Section V summarizes our findings and concludes the paper.

III. A basic mixed integer programming formulation

We consider the GMSP for hydropower systems in the general form

$$\max_{\substack{y \in \mathcal{Y} \\ x(y) \in \mathcal{X}(y)}} \quad \Phi(x(y)) - \Psi(y), \tag{2}$$

where the variables y denote the maintenance decisions and the variables $x(y)$ represent operational decisions, such as turbine discharges and water spills. The feasible set $\mathcal{X}(y)$ of the operational decisions is determined by the water balance constraints and the bounds of the hydropower operation, which are affected by the scheduled outages y. The set $\mathcal{Y}$ of feasible maintenance decisions is defined by the time window constraints of the maintenance activities, the maximum number of simultaneous maintenance outages, and other logical constraints. In (2), the functions $\Psi(y)$ and $\Phi(x(y))$ denote the maintenance cost, the value of the electricity production during the planning horizon, and the value of the stored water at the end of the planning horizon, respectively. Note that the value of the electricity production $\Phi(x(y))$ is affected by the maintenance schedule y because the power production function is different for each set of active generators (Fig. 1). The interdependency between the maintenance plan and the hydropower operation makes this a challenging nonlinear, nonconvex and combinatorial optimization problem.

In the next subsections we formulate in turn the hydropower operation, the linear approximation of the power production function, and the maintenance scheduling.

A. The hydropower operation

The hydropower operation problem optimizes the water discharges, water spills and stored water levels to maximize the total expected value of the electricity production, while respecting the physical constraints of the system and the target levels of the reservoirs at the end of the planning horizon. The physical constraints enforce the mass and power balance, as well as the bounds of the variables, such as the water levels in reservoirs. At each time period $t \in \mathcal{T}$, reservoirs can be fed by lateral inflows F_{it} from tributary rivers or snow-melt, or by turbine discharges u_{gt} and water spills v_{gt} from upstream reservoirs $g \in \mathcal{U}(i)$.

At each powerhouse and time period, the mass balance (3) implies that the initial water volume $s_{i(t-1)}$ minus the water volume s_{it} at the end of the time period should be equal to the water inflows minus the total outflows, multiplied by the conversion factor Q. As it is customary, we assume that the outflows are equal to the total turbine discharge u_{it} and the water spill v_{it} of the reservoir.

$$s_{it} - s_{i(t-1)} = Q\left(F_{it} + \sum_{g \in \mathcal{U}(i)} [u_{gt} + v_{gt}] - u_{it} - v_{it}\right), \quad \forall\, t \in \mathcal{T},\, i \in \mathcal{I}. \tag{3}$$

To ensure the consistency with the initial and the final water volume of the reservoirs, we define $s_{i(t-1)} = S_{i0}$ for $t = 1$ in (3). In addition, (4)-(6) define the bounds on the water discharge, water spill and water volume.

$$0 \le u_{it} \le \bar{U}_i, \quad \forall i \in \mathcal{I},\, t \in \mathcal{T}, \tag{4}$$

$$0 \le v_{it} \le \bar{V}_i, \quad \forall i \in \mathcal{I},\, t \in \mathcal{T}, \tag{5}$$

$$\underline{S}_i \le s_{it} \le \bar{S}_i, \quad \forall\; i \in \mathcal{I},\, t \in \mathcal{T}. \tag{6}$$

The energy balance (7) requires that at each time period t, the total energy production plus the energy purchases equal the load J_t plus the energy sales:

$$\sum_{i \in \mathcal{I}} p_{it} + w_t^- = J_t + w_t^+, \quad \forall\, t \in \mathcal{T}, \tag{7}$$

with bounded electricity trade variables,

$$0 \le w_t^+ \le \bar{W}^+, \quad \forall\, t \in \mathcal{T}, \tag{8}$$

$$0 \le w_t^- \le \bar{W}^-, \quad \forall\, t \in \mathcal{T}. \tag{9}$$

Notice that this definition of the power balance in (7) can describe a variety of situations for electricity producers. For example, the parameter $J_t > 0$ can represent the case of a producer that in a liberalized electricity market has committed to supply an amount of electricity J_t in the forward market, and that in the spot or day ahead market can trade electricity (w_t^-, w_t^+) to compensate for the differences between its forward commitment J_t and its actual electricity production. Clearly, if at some time period $t \in \mathcal{T}$ the electricity purchase is not allowed, it suffices to define $w_t^- = 0$.

B. Linearization of the power production function

As the nonlinearity of the electricity production functions (Fig. 1) poses a challenge to the solution of the GMSP, we approximate these functions with linear inequalities. In this way, we can formulate the GMSP as a mixed-integer linear program (MILP), which can be tackled with state-of-the art solvers [13].
For each powerhouse, the power output p_{it} is a nonlinear function Θ_i of the water discharge u_{it} and the net water head (which in turn is a nonlinear function of the stored water volume s_{it} and the total water discharge u_{it}). Since each generator may have a particular efficiency curve, the maximum power output in a powerhouse depends on the specific set of active generators. However, if the differences among power functions of individual generators are negligible, the power function in a powerhouse can be characterized by the number of active generators k_{it}, instead of the explicit set of active generators, that is, $p_{it} = \Theta_i(s_{it}, u_{it}, k_{it})$. This assumption significantly reduces the problem complexity, since otherwise a specific power function would be necessary for each combination of active generators.

For each number of active generators with their respective efficiency curves, a dynamic programming algorithm can determine the commitment of units, as well as the maximum power output corresponding to a set of water discharges and stored water levels (Fig. 1) [12]. By definition, this set of points is contained in its convex hull, whose half-space representation can be obtained with a facet enumeration algorithm. Some implementations of this algorithm are freely available [14], [15].

The resulting polyhedron may contain a large number of hyperplanes, some of which should be dropped since they define the lower facets of the convex hull with respect to the power output p_{it}. The set can be additionally reduced by iteratively removing the hyperplane for which the remaining polyhedron has the smallest approximation error of the power output. This sequential elimination of hyperplanes is repeated until the target number of hyperplanes or a specified precision is reached.

For each powerhouse i and number of active generators k, the resulting set of hyperplanes $\mathcal{H}(i,k)$ with parameters β^0_h, β^u_h and β^s_h provides an outer approximation $\hat{p}_{it}$ of the power output corresponding to the specific amounts of water discharge u_{it} and stored water level s_{it}, when k generators are active, i.e.,

$$\begin{gathered} 0 \le \hat{p}_{it} \le \beta^0_h + \beta^u_h u_{it} + \beta^s_h s_{it}\,, \\ \forall\ i \in \mathcal{I},\, t \in \mathcal{T},\, k \in \mathcal{K}(i,t),\, h \in \mathcal{H}(i,k). \end{gathered}$$

Notice that through the index $h \in \mathcal{H}(i,k)$, the hyperplane parameters β^0_h, β^u_h and β^s_h are defined for the corresponding powerhouse i and number of active generators k.
At powerhouse i and time period t, if k^* is the number of active generators, the power function constraints for $k \neq k^*$ can be relaxed by adding the bounding term $(1 - z_{itk})\bar{P}_i$ on the right hand side of (10), i.e.,

$$\begin{gathered} 0 \le \hat{p}_{it} \le \beta^0_h + \beta^u_h u_{it} + \beta^s_h s_{it} + (1 - z_{itk})\bar{P}_i, \\ \forall\ i \in \mathcal{I},\, t \in \mathcal{T},\, k \in \mathcal{K}(i,t),\, h \in \mathcal{H}(i,k), \end{gathered} \tag{10}$$

where $\bar{P}_i$ is the generation capacity of powerhouse i and the binary variables z_{itk} indicate if k generators are active at (i,t). Since only one binary variable z_{itk} takes value 1 for each $(i,t) \in \mathcal{I} \times \mathcal{T}$,

$$\sum_{k \in \mathcal{K}(i,t)} z_{itk} = 1,\ \forall\ i \in \mathcal{I}, t \in \mathcal{T}. \tag{11}$$

Thus, by (11) and the binary condition on z_{itk}, the approximated power output $\hat{p}_{it}$ in (10) is bounded only by the hyper-plane set corresponding to the number of active turbines.
The quality of the approximation given by (10) increases with the number of hyperplanes in $\mathcal{H}(i,k)$ and with the convexity of the actual power production function. Thus, there is a compromise between model size and solution quality. In our tests with real data the approximation errors of this approach were 0.5% and 0.25% of the electricity production for 15 and 30 hyperplanes in $\mathcal{H}(i,k)$, respectively. Nevertheless, the overestimate of the power production can be reduced with

$$p_{it} = \alpha_0 + \alpha_1 \hat{p}_{it},\ \forall\ i \in \mathcal{I}, t \in \mathcal{T}, \tag{12}$$

where p_{it} is the corrected estimate of the electricity production and α_0, α_1 are the parameters of a linear regression model that fits the estimated electricity production $\hat{p}_{it}$ to the corresponding amounts of actual electricity production, using historical data.

C. The maintenance scheduling problem

For each maintenance activity $m \in \mathcal{M}$, the interval between the earliest starting time period E_m and the latest starting time period L_m defines the set of time periods $\mathcal{T}(m)$ when the activity m can start: $\mathcal{T}(m) = \{t \in \mathcal{T} \,|\, E_m \le t \le L_m\}$. We assume that each activity can be completed within the planning horizon $\mathcal{T}$, i.e., $E_m \le L_m \le T - D_m + 1$, where D_m denotes the duration of the maintenance task m.

The definition of the binary variables $y_{mt}, \forall\, m \in \mathcal{M}, t \in \mathcal{T}(m)$, for representing the maintenance decisions (see notation in Section I) avoids the definition of time window constraints since the set $\mathcal{T}(m)$ encodes the time window parameters of each activity. Unnecessary y_{mt} variables are excluded from the model because they are defined using $\mathcal{T}(m)$ instead of $\mathcal{T}$.

For the basic maintenance problem we consider only the constraints on: completion of maintenance tasks, maximum number of generator outages, and mapping the maintenance schedule to the number of active generators.

The task completion constraints (13) enforce each activity to start at one of the feasible time periods $\mathcal{T}(m)$. Constraints (14) compute for each powerhouse the number of maintenance activities r_{it} in execution at time period t, among the set of activities $\mathcal{M}(i)$ that must be completed at station i.

$$\sum_{t\in\mathcal{T}(m)} y_{mt} = 1, \quad \forall\, m \in \mathcal{M}. \tag{13}$$

$$\sum_{\substack{m\in\mathcal{M}(i) \\ t'\in\{\mathcal{T}(m)\,|\,(t-D_m+1)\le t'\le t\}}} y_{mt'} = r_{it}, \quad \forall\, i \in \mathcal{I}, t \in \mathcal{T}. \tag{14}$$

Notice that at time period t an activity m is in execution if it starts between $t - D_m + 1$ and t. This is the interval of index t' on the summation term in (14).

The maximum number of outages O_{it} bounds r_{it}:

$$0 \le r_{it} \le O_{it}, \quad \forall\, i \in \mathcal{I}, t \in \mathcal{T}. \tag{15}$$

O_{it} depends on the maintenance resources. In addition, for a feasible operation, O_{it} cannot exceed the difference between the number of available generators $\bar{G}_{it}$ and the minimum number of generators in service $\underline{G}_i$, i.e., $O_{it} \le \bar{G}_{it} - \underline{G}_i$, $\forall\, i \in \mathcal{I}, t \in \mathcal{T}$. Notice that $\bar{G}_{it}$ is a time varying parameter, since the number of available generators can be affected by existing generator outages or by previous maintenance scheduling decisions. On the other hand, the minimum number of generators $\underline{G}_i$ is constant in time due to operational requirements.

Constraints (16) map the number of outages r_{it} into the variables z_{itk}. At each period and powerhouse, the maximum number of available generators $\bar{G}_{it}$ equals the sum of the number of outages r_{it} plus the number of active generators k^* corresponding to $z_{itk^*} = 1$.

$$r_{it} + \sum_{k\in\mathcal{K}(i,t)} k z_{itk} = \bar{G}_{it}, \forall\, i \in \mathcal{I}, t \in \mathcal{T}. \tag{16}$$

Constraints (18)-(17) specify the binary decision variables.

$$z_{itk} \in \{0, 1\}, \forall\, i \in \mathcal{I}, t \in \mathcal{T},\, k \in \mathcal{K}(i,t), \tag{17}$$

$$y_{mt} \in \{0, 1\}, \forall m \in \mathcal{M},\, i \in \mathcal{T}(m). \tag{18}$$

D. The objective function

The GMSP maximizes the value of the electricity production plus the value of the stored water, minus the sum of the maintenance costs:

$$\underset{\substack{w^+,w^-,u,v,s,\\ p,y,z}}{\text{Maximize}} \; \sum_{t\in\mathcal{T}} \left(B_t^+ w_t^+ - B_t^- w_t^-\right) - \sum_{\substack{m\in\mathcal{M},\\ t\in\mathcal{T}(m)}} C_{mt} y_{mt}, \tag{19}$$

The value of the electricity production during the planning horizon is calculated as the net benefit of the electricity trade, i.e., the difference between the revenue of electricity sale ($B_t^+ w_t^+$) and the cost of electricity purchase ($B_t^- w_t^-$).

E. The complete basic model

We refer to the resulting mixed-integer linear programming (MILP) problem as P_B:

Maximize (19) subject to constraints (3) - (18).

Notice that for any feasible maintenance schedule $(\bar{y}, \bar{z})$, the resulting hydropower operation subproblem P_H is the linear program

$$P_H(\bar{y}, \bar{z}) = \underset{w^+,w^-,u,v,s,p}{\text{Maximize}} \; \sum_{t\in\mathcal{T}} \left(B_t^+ w_t^+ - B_t^- w_t^-\right), \tag{20}$$

subject to (3) -(10), (12).

Naturally, in P_H, the simultaneous purchase and sale of electricity (i.e., the case of arbitrage) can be prevented if the sale price of the electricity B_t^+ is lower than the purchase price B_t^- as stated in Proposition 1.

Proposition 1. *In any optimal solution to $P_H(\bar{y}, \bar{z})$ with electricity prices $B_t^+ < B_t^-$, either $w_t^+ = 0$ or $w_t^- = 0$.*

See Appendix A for a proof of this proposition.

Furthermore, this property also holds for any feasible solution to P_B obtained with a general MILP solver (e.g., CPLEX Gurobi, Xpress-MP), even if the maintenance schedule is not optimal. Indeed, any feasible solution returned by such a solver is obtained at a node of the search tree by solving a linear program to optimality.

IV. Tightening approaches

Due to the weak continuous relaxation of (10) and (16), the formulation in Section III-E is difficult to solve for realistic instances. In this section we explore three approaches for tightening the formulation: extended formulation, set reduction and valid inequalities.

A. Extended formulation

The bound (10) can be very weak because it is valid for any operating condition and for any number of active generators k on the interval $(\bar{G}_{it}, \underline{G}_i)$. However, $\bar{P}_{ik}$ and p_{itk} can be based on the actual number of active generators k and the specific operating conditions at each time period and powerhouse. Constraints (21) specify the power bound for each number of active generators, and (22) ensure the equivalence with the original variables p_{it} and in substitution of (10), constraints (23) define a linear approximation of the power function.

$$0 \le p_{itk} \le z_{ik} \bar{P}_{ik}, \quad \forall\, i \in \mathcal{I},\, t \in \mathcal{T},\, k \in \mathcal{K}(i,t). \tag{21}$$

$$\sum_{k\in\mathcal{K}(i,t)} p_{itk} = p_{it}, \quad \forall\, i \in \mathcal{I},\, t \in \mathcal{T}, \tag{22}$$

$$\begin{gathered} 0 \le p_{itk} \le \beta_h^0 + \beta_h^u u_{it} + \beta_h^s s_{it}, \\ \forall\, i \in \mathcal{I},\, t \in \mathcal{T},\, k \in \mathcal{K}(i,t), h \in \mathcal{H}(i,k), \end{gathered} \tag{23}$$

Thus we have P_E as the MILP with the extended formulation:

Maximize (19) subject to (3)-(9), (11)-(18), (21)-(23).

The bounds $\bar{P}_{ik}$ for (21) can be obtained as the optimal values q_{ik}^* from maximizing the power output in (23) when the stored water level is maximum:

$$\underset{q,u}{\text{Maximize}}\ q_{ik} \text{ s.t. } q_{ik} \leq \beta_h^0 + \beta_h^u u_{itk} + \beta_h^s \bar{S}_i,\ \forall\, h \in \mathcal{H}(i,k). \tag{24}$$

B. Set reduction

Next we exploit the time window parameters of the maintenance tasks in order to exclude unnecessary set elements. As a consequence, fewer constraints and variables are defined, leading to a tighter continuous relaxation and fewer choices for branching. We aim at reducing the set $\mathcal{K}(i,t)$ that determines both the number of binary variables z_{itk} and the degrees of freedom of the system (11) and (16). A maintenance activity m beginning at E_m and with duration D_m spans the interval $\mathcal{T}^E(m) = \{\, t \in \mathcal{T}(m) \,|\, E_m \leq\, t < E_m + D_m\}$. Likewise, if activity m starts at L_m, it spans the interval $\mathcal{T}^L(m) = \{\, t \in \mathcal{T}(m) \,|\, L_m \leq t < L_m + D_m\}$. The overlap of the two intervals

$$\begin{aligned}\mathcal{T}^{\mathcal{O}}(m) &\triangleq \mathcal{T}^E(m) \cap \mathcal{T}^L(m) \\ &= \{\, t \in \mathcal{T}(m) \,|\, L_m\ \leq t\ \leq\ E_m + D_m\},\end{aligned}$$

defines the set of time periods when the activity necessarily will take place. Likewise, the span of a maintenance activity m is the interval $\mathcal{T}^{\mathcal{S}}(m)$ where the activity can be in execution. Since activity m cannot start before E_m and it can finish no later than $L_m + D_m$, we define

$$\mathcal{T}^{\mathcal{S}}(m) = \{\, t \in \mathcal{T}(m) \,|\, E_m\ \leq t\ \leq\ L_m + D_m\}.$$

These definitions are illustrated in Fig. 2.

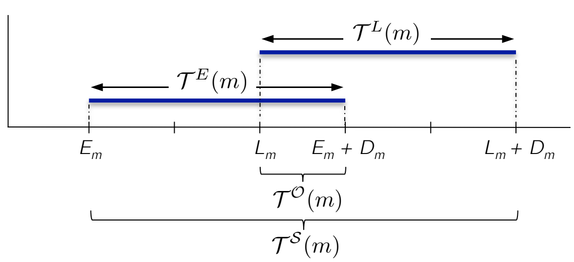


Fig. 2: Timeline for a maintenance activity m.

The maximum number of maintenance activities that *can* be in execution at powerhouse i during time period t is the cardinality of the set of tasks whose spans $\mathcal{T}^{\mathcal{S}}(m)$ intersect at time period t, that is,

$$\bar{R}_{it} = |\{\, m \in \mathcal{M}(i) \,|\, t \in \mathcal{T}^{\mathcal{S}}(m)\,\}|.$$

Similarly, the minimum number of activities in execution at powerhouse i during time period t is,

$$\underline{R}_{it} = |\{\, m \in \mathcal{M}(i) \,|\, t \in \mathcal{T}^{\mathcal{O}}(m)\,\}|.$$

Naturally, $\bar{R}_{it}$ and $\underline{R}_{it}$ bound the number of outages r_{it}:

$$\underline{R}_{it} \leq r_{it} \leq\ \bar{R}_{it},\ \ \forall\ i \in \mathcal{I}, t \in \mathcal{T}. \tag{25}$$

Proposition 2. *In formulations P_B and P_E, the feasible number of active generators k at period $t \in \mathcal{T}$ and powerhouse $i \in \mathcal{I}$ is in the set*

$$\mathcal{K}(i,t) = \{\ k\ \in \mathbb{Z} \,|\ \underline{K}_{it}\ \leq\ k\ \leq\ \bar{K}_{it}\},\tag{26}$$

where

$$\underline{K}_{it} = \max\{\bar{G}_{it} - O_{it},\, \bar{G}_{it} - \bar{R}_{it}\}, \tag{27}$$

$$\bar{K}_{it} = \bar{G}_{it} - \underline{R}_{it}. \tag{28}$$

See Appendix B for a proof of this proposition.

From (26-28) we see that the greater the difference between $\bar{G}_{it}$ and $\bar{K}_{it}$, as well as between $\underline{G}_i$ and $\underline{K}_{it}$, the greater the reduction in the number of variables and constraints with index $k \in \mathcal{K}(i,t)$.

C. Valid inequalities

Finally, we analyze the linear system formed by constraints (11) and (16), which in general is undetermined and has multiple non-integer solutions. We consider the case when $\underline{R}_{it} = 0$.

If $r_{it} = 0$, then from constraints (16), $\sum_{k \in \mathcal{K}(i,t)} z_{itk} k = \bar{G}_{it}$, which implies $z_{itk} = 1$ for $k = \bar{G}_{it}$, since by constraint (11) only one binary variable z_{itk} should be active for each $(i,t) \in \mathcal{I} \times \mathcal{T}$. On the other hand, if $r_{it} \geq 1$, then $z_{itk} = 0$ for $k = \bar{G}_{it}$ with $(i,t) \in \mathcal{I} \times \mathcal{T}$. By disaggregating r_{it} into the corresponding y_{mt} variables (see (14)), these logical implications are equivalent to

$$\sum_{t' \,\in\, \{\mathcal{T}(m) \,|\, (t - D_m + 1)\, \leq t'\, \leq t\}} y_{mt'} + z_{itk} \leq 1, \quad \text{for } k = \bar{G}_{it}, \\ \forall\ i \in \mathcal{I},\, m \in \mathcal{M}(i),\, t \in \mathcal{T}, \tag{29}$$

which by the binary condition on z_{itk} and y_{mt} are facet defining inequalities.

Also, $r_{it} = 0$ implies $z_{itk} = 0 \;\forall\; k \in \{\mathcal{K}(i,t) \setminus \bar{G}_{it}\}$:

$$\sum_{k \,\in\, \mathcal{K}(i,t) \,\setminus\, \bar{G}_{it}} z_{itk}\ \leq r_{it},\ \ \forall\ i \in \mathcal{I},\, t \in \mathcal{T}. \tag{30}$$

Next we show that constraints (29)-(30) allow relaxing the integrality of a subset of binary variables z_{itk} when $\bar{K}_{it} = \bar{G}_{it}$ and the number of degrees of freedom of the system (11),(16) is sufficiently small.

Proposition 3. *In models P_B and P_E with constraints* (29)-(30) *if for some $(i', t') \in \mathcal{I} \times \mathcal{T}$:*

i) $\underline{R}_{i't'} = 0$,
ii) $\bar{K}_{i't'} - \underline{K}_{i't'} \leq 2$,
iii) there exists an integer feasible solution,

then the integrality condition (17) *for $z_{i't'k}\ \forall\, k \in \mathcal{K}(i', t')$ can be relaxed as the variables $z_{i't'k}$ will be integer in any feasible solution.*

See Appendix C for a proof of this proposition.

V. Computational experiments

In this section we report on our computational experiments to evaluate the eight formulations obtained starting from the basic model and including/excluding each of the three approaches in Section IV. The eight combinations are given in Table II, where 1 indicates that a given approach is used in the corresponding formulation, and 0 indicates that the approach was not used.

We conducted two experiments to determine the best combination. First, we solved smaller instances of GMSP and analyzed the computation times in order to select a subset of formulations. Second, we evaluated this subset via experiments with larger instances.

Our test instances were adapted from a cascade 4-powerhouse system. For each powerhouse and number of generators, we approximated the hydropower production function with 30 linear inequality constraints (10) and (23). For each instance, maintenance requirements are specified with the following parameters for each activity: index, powerhouse, duration, earliest start time period, and latest start time period. We maximize the value of the electricity production, with a sale price of 8 \$/kWh, and $d_t = 0$ and $w_t^- = 0, \forall t \in \mathcal{T}$.

A. Computational results for all formulations

For the first experiment, we defined two levels for each of the five factors of the instance size (Table I). For each

TABLE I: Levels of factors used to create the test intances to compare all formulations in Section V-A.

Factor	Low Level	High Level
Number of maintenance tasks	8	10
Number of time periods	20	25
Time window length	5	8
Maximum outages in each powerhouse	2	3
Avg. duration of maintenance tasks	4	5

of the $2^5 = 32$ combinations of these factors, we created two maintenance datasets, for a total of 64 test instances. The size of the MILP formulations ranged from 94 binary variables, 390 continuous variables and 4263 constraints, to 456 binary variables, 775 continuous variables and 12485 constraints. Because randomly generating instances for GMSP is prone to infeasibilities, we created new instances with random changes in a subset of parameters of initial feasible instances. When an infeasible instance was obtained by this procedure, we restored its feasibility by arbitrarily changing the instance parameters.

We ran the tests in a 24-processor Intel® Xeon® server at 2.7 GHz with 32.9 GB RAM, with 4 cores dedicated for running the Xpress-MP solver. The models were coded in C++ with the Xpress BCL 8.1.0 callable library [16].

We chose CPU clock time as the basic performance metric, which allows to measure the actual computation time for solving the problem, without the effect of background processes. Given that the computation times increase significantly with the size of the instance and also differ between instances of similar size, we normalized for each instance the logarithmic CPU time according to the standard score

$$z_{jb} = (t_{jb} - \mu_j^t)/\sigma_j^t, \tag{31}$$

where t_{jb} is the logarithmic CPU time for solving instance $j \in \mathcal{J}$ with formulation $b \in \mathcal{B}$, and μ_j^t, σ_j^t are respectively the mean and the standard deviation of the logarithmic CPU times of the 8 models for solving instance j.

We report in Table II the mean $\bar{z}_b$ and standard deviation σ_b^z of z_{jb} over the 64 test instances, for each formulation. The results show that the choice of formulation affects the computation times, as corroborated with a p-value of 0.005 for a one-way ANOVA, which for a significance level of $\alpha = 0.01$ indicates a significant effect of the selected formulation on the logarithmic CPU time.

TABLE II: Normalized log CPU times per instance, computed from 64 test instances.

Formu-	Tightening approaches			Norm. log CPU time	
lation	Set reduc.	Valid ineq.	Extended formul.	Average $\bar{z}_b$	St. dev. σ_b^z
1	0	0	0	**1.469**	0.35
2	0	0	1	-0.849	0.40
3	0	1	0	0.790	0.38
4	0	1	1	-0.685	0.33
5	1	0	0	0.421	**0.50**
6	**1**	**0**	**1**	**-0.880**	**0.34**
7	1	1	0	0.511	0.39
8	1	1	1	-0.776	0.42

In these instances, the wall-clock time to reach optimality ranged from 1 s to 1743 s, with an average of 84.27 s over all formulations. The computational wall-clock time was highly correlated with the CPU time ($R^2 = 0.99$).

While formulation 1 had the largest average normalized log CPU time, the smallest time was achieved by formulation 6 (extended formulation with set reduction). The latter also had the second smallest standard deviation. The maximum standard deviation corresponded to the formulation with only set reduction. Overall, the formulations 2, 4, 6, and 8 gave the best results in Table II. In several instances, we registered more than one order of magnitude of difference in wall-clock time between the basic model (formulation 1) and the best formulation. However, in Table II these differences are attenuated by the logarithmic transformation that we applied.

The effect of the choice of formulation also shows in the performance profiles of Fig. 3. A performance profile [17] gives the cumulative relative frequency $\rho_b(\tau)$ with which a formulation solves instances of the problem within a factor τ of the best possible value of $\log_2(r_{jb})$, where $r_{jb} = t_{jb}/\min_{b \in B} t_{jb}$, and

$$\rho_b(\tau) = \frac{1}{n_j} \operatorname{size}\{j \in \mathcal{J} : \log_2(r_{jb}) \le \tau\}. \tag{32}$$

In summary, the curves closest to the top left corner correspond to the formulation with the best performance.

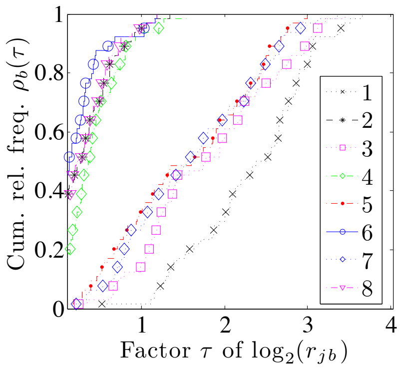


Fig. 3: Performance profiles of the tested formulations

Fig. 3 shows that the formulations with at least one tightening component perform better than the basic model (formulation 1). In Fig. 3, the performance profiles of the best 4 formulations indicate that formulation 6 is a clear winner for $\tau \leq 0.8$. In less than 10% of the instances, models 2 and 8 are a competitive choice.

The extended formulation is common to the 4 best-performing formulations in Fig. 3. The ANOVA results in Table III show that this approach, either alone or in combination with others, has a significant effect for arbitrarily small significance α levels (p-value = 2.36e-12). On the other hand, although the formulation with only valid inequalities outperformed the basic model (formulation 3 vs. formulation 1 in Table II and Fig. 3), the effect of the valid inequalities was not statistically significant when combined with other formulation approaches (p-value = 0.758). Finally, the effect of set reduction is only significant for $\alpha \geq 0.2$ (p-value = 0.181).

TABLE III: p-values based on normalized log CPU time

Approach	p-value
Set reduction	0.181
Valid inequalities	0.758
Extended formulation	2.36e-12

B. Optimality gaps of the best formulations

In the second experiment, we worked only with formulations 2, 6 and 8. These have the smallest average CPU times in Table II, and clearly outperform formulation 4 in Fig. 3. Our focus is on the optimality gaps that these formulations can achieve for large instances of GMSP. We tested these formulations on 16 instances with more maintenance tasks than the earlier instances. These 16 instances were generated with two maintenance datasets for each of the $2^3 = 8$ combinations of the levels of the three factors in Table IV. For these instances we specified a planning horizon with 25 time periods in a cascade 4-powerhouse system, with a maximum of 2 outages in each powerhouse.

Table V reports the optimality gap statistics for the three formulations after 1,000 and 20,000 seconds of CPU time on each instance.

TABLE IV: Levels of factors for the test instances to compare the best formulations V-B.

Factor	Low Level	High Level
Number of maintenance tasks	15	20
Time window length	5	8
Avg. duration of maintenance tasks	4	5

TABLE V: Optimality gap statistics

Formulation	CPU time 20,000 s		CPU time 1,000 s	
	Mean	St. dev.	Mean	St. Dev
2	0.0144	0.0069	0.0295	0.0235
6	0.0144	0.0071	0.0229	0.0076
8	0.0151	0.0073	0.0273	0.0222

All three formulations reached average optimality gaps below 3 % within 1,000 CPU s. Progress is substantially slower after that, and at the time limit of 20,000 CPU s the average optimality gap in all three formulations is close to 1.5%. Formulation 6 had the best overall performance after 1,000 CPU s, and formulations 2 and 6 had similar average performance after 20,000 CPU s. The average wall-clock time corresponding to the CPU time limit of 20,000 CPU s was 2,955.5 s, with a standard deviation of 51 s. Due to the specified time limit in this experiment, the optimal solution was not reached in any of the runs. However, the small optimality gaps in Table V indicate that with computational times beyond the specified time limit, the optimal solutions for the instance sizes that we considered are achievable in practise.

Based on the overall results, we conclude that the most promising approach is the extended formulation with set reduction (formulation 6), and possibly in combination with the valid inequalities.

VI. Industrial application

We tested this approach with data adapted from a 4-powerhouse system of Rio Tinto in the Saguenay-Lac-St-Jean region in Québec, Canada (see Table VI). At the company, turbine-generator systems must undertake periodic preventive maintenance tasks of short duration. Less frequently, activities of longer duration, such as overhauling of generators, are also necessary. We considered 18 maintenance tasks to be completed in a planning horizon of 30 days. For each task, the time window, as well as the starting time of the activity according of an initial maintenance schedule are given. As in the previous section, the electricity production for each number of generators and powerhouse was approximated with 30 hyperplanes, and we set $d_t = 0$ and $w_t^- = 0$, $\forall t \in \mathcal{T}$. For this application, the relevant price is 5 ¢/kWh.

We used formulation 6 (with set reduction and valid inequalities) to solve this instance of the problem with the Xpress-MP solver in deterministic mode with 20 threads in a 24-processor Intel® Xeon® server at 2.7 GHz with 32.9 GB RAM. As previous works on the GMSP [2]-[5] did not consider the maintenance time windows and other

TABLE VI: Basic attributes of the hydropower system. Powerhouses are ordered from upstream to downstream.

System type	Number of generators	Installed capacity (MW)	Maintenance tasks
Reservoir	5	205	4
Run of the river	5	210	5
Reservoir	12	402	4
Run of the river	17	1587	5
Total	39	2404	18

relevant aspects of the problem, these approaches can lead to infeasible solutions in practice. For this reason, in this industrial application example, we compare the value of the solution obtained with our model against the optimal maintenance schedule obtained with a simplified model P_S that neglects the nonlinearity of the electricity production, while still respecting the time windows of the maintenance tasks. Thus, we relax (23) to define P_S as

Maximize (19) subject to (3)-(9), (11)-(18), (21)-(22).

For the application example in this Section, the proposed model (formulation 6) has 7103 continuous variables, 299 binary variables and 7402 constraints. After 1207 s an optimal solution was found with an objective value of \$ 57.802 M. In contrast, the best solution found with the simplified model (P_S) has an objective value of \$ 67.444 M. The higher objective value of this solution is merely a consequence of the overestimated electricity production in P_S by ignoring the nonlinearity of the HPF. When the actual nonlinearity of the hydroelectricity production is considered, the maximum revenue of the maintenance schedule obtained with P_S is \$ 57.735 M. With respect to this solution, the optimal schedule of formulation 6 yields an increase of 1340 MWh of electricity production in the one-month planning horizon and an approximate annualized gain of \$ 804,000. The increment of the electricity production in the optimal solution is mainly a consequence of the reduction of accumulated water spills during the planning horizon, which translates into higher average stored water level and more efficient operation of the generators.

VII. Conclusions

We proposed a mixed-integer optimization model for the GMSP in hydropower systems, and three possible approaches to tighten its continuous relaxation: set reduction, valid inequalities, and extended formulation. Using a set of 64 test instances, we found that the extended formulation had the most significant effect in decreasing the computational time, and that the combination of extended formulation and set reduction achieved the best average performance and small variability in computation time. This formulation was tested in a real 4-powerhouse hydropower system with 39 generators and 2404 MW of generation capacity, and an optimal maintenance schedule for a one-month planning horizon was found in less than 30 minutes.

We proved that under some conditions, the valid inequalities allow relaxing the integrality condition on a subset of binary variables of the problem. Although this insight did not exhibit a statistically significant effect in our tests, we consider that the mathematical result can be useful for developing heuristic solution methods for this problem as well as for other problems with similar integer-mapping constraints.

Because the GMSP typically spans a planning horizon of several weeks, in practice it may be possible to run the solver for several hours or even days, in order to obtain either optimal or near optimal solutions. However, more efficient solution methods are necessary to solve larger real instances. Furthermore, incorporating other relevant aspects of the problem, such as transmission system effects and uncertainty of water inflows will increase the computational complexity of the problem. Solution approaches considering these elements will be the subject of future work.

Appendix A
Proof of Proposition 1

Proof. By contradiction, suppose that $w_t^+ > 0$ and $w_t^- > 0$ for some t. Consider 3 cases: i) $w_t^+ > w_t^-$, ii) $w_t^- > w_t^+$, and iii) $w_t^- = w_t^+$. In case i), $B_t^- > B_t^+$ implies $-B_t^- w_t^- < -B_t^+ w_t^-$. Adding $B_t^+ w_t^+$ gives $B_t^+ w_t^+ - B_t^- w_t^- < B_t^+ w_t^+ - B_t^+ w_t^- = B_t^+(w_t^+ - w_t^-) = B_t^+ q_t^+$, which shows that $w_t^+ > w_t^- > 0$ is not optimal, since selling $q_t^+ = w_t^+ - w_t^-$ reaches higher profit than buying w_t^- and selling w_t^+. In case ii), $B_t^- > B_t^+$ implies $B_t^+ w_t^+ < B_t^- w_t^+$. Subtracting $B_t^- w_t^-$ gives $B_t^+ w_t^+ - B_t^- w_t^- < B_t^- w_t^+ - B_t^- w_t^- = B_t^-(w_t^+ - w_t^-) = -B_t^- q_t^-$, which shows that $w_t^- > w_t^+ > 0$ is not optimal, since the net cost of buying w_t^- and selling w_t^+ is higher than the cost of selling $q_t^- = w_t^+ - w_t^-$. In case iii), $w_t^- = w_t^+ = w_t$ implies $w_t(B_t^+ - B_t^-)$. As $B_t^- > B_t^+$, $w_t = 0$ minimizes the loss. Since w_t^- and w_t^+ cannot be both positive in any case, either $w_t^- = 0$ or $w_t^+ = 0, \forall\, t \in \mathcal{T}$ in any optimal solution. □

Appendix B
Proof of Proposition 2

Proof. From (15) and (25),

$$r_{it} \leq \min\{O_{it},\, \bar{R}_{it}\}\,, \quad \forall\, i \in \mathcal{I}, t \in \mathcal{T}. \tag{33}$$

From (16),

$$\begin{aligned}
\sum_{k \in \mathcal{K}(i,t)} k z_{itk} &= \bar{G}_{it} - r_{it}, \forall\, i \in \mathcal{I}, t \in \mathcal{T}, \\
&\geq \bar{G}_{it} - \max\{r_{it}\}, \forall\, i \in \mathcal{I}, t \in \mathcal{T}, \\
&= \bar{G}_{it} - \min\{O_{it},\, \bar{R}_{it}\}, \forall\, i \in \mathcal{I}, t \in \mathcal{T}, \\
&\qquad \text{(by Eq. 33)} \\
&= \max\{\bar{G}_{it} - O_{it},\, \bar{G}_{it} - \bar{R}_{it}\}, \forall\, i \in \mathcal{I}, t \in \mathcal{T}, \\
&\triangleq \underline{K}_{it}.
\end{aligned}$$

Then, by (11) and (16), $k \geq \underline{K}_{it}, \forall k \in \mathcal{K}(i,t)$. Similarly, from constraints (16),

$$\begin{aligned}\sum_{k\in\mathcal{K}(i,t)} kz_{itk} &= \bar{G}_{it} - r_{it}, \forall i \in \mathcal{I}, t \in \mathcal{T},\\ &\leq \bar{G}_{it} - \min\{r_{it}\}, \forall i \in \mathcal{I}, t \in \mathcal{T},\\ &= \bar{G}_{it} - \underline{R}_{it}, \forall i \in \mathcal{I}, t \in \mathcal{T},\\ &\triangleq \bar{K}_{it},\end{aligned}$$

which by (11) and (16) implies $k \leq \bar{K}_{it}, \forall k \in \mathcal{K}(i,t)$. □

Appendix C
Proof of Proposition 3

Proof. To simplify the notation, we drop the indices $(i',t') \in \mathcal{I} \times \mathcal{T}$ from $\bar{K}_{i't'}$, $\bar{R}_{i't'}$, $\underline{R}_{i't'}$, $r_{i't'}$, $\mathcal{K}(i',t')$ and $z_{i't'k}$. In any feasible solution to P_B, P_E, variables y_{mt} are binary by (18) and r is integer by (14). By condition i), all available $\bar{G}$ generators can be active, which implies $\bar{K} = \bar{G}$ according to (28). Condition i) also implies $r \geq 0$ by (25). On the other hand, by (16) and Condition ii), $r \leq 2 = \bar{R}$. Therefore, for the analysis of the linear system with (11) and (16), we consider three cases:

1) $r = 0$: By conditions i) and ii),

$$\mathcal{K} = \{\bar{G}, \bar{G}-1, \bar{G}-2\}. \tag{34}$$

Then, the linear system (11) and (16) can be written in extensive form as

$$z_{\bar{G}} + z_{\bar{G}-1} + z_{\bar{G}-2} = 1, \tag{35}$$

$$\bar{G}z_{\bar{G}} + (\bar{G}-1)z_{\bar{G}-1} + (\bar{G}-2)z_{\bar{G}-2} = \bar{G} - r. \tag{36}$$

By (30), $r = 0$ implies $z_k = 0 \; \forall k < \bar{G}$. Then, by (11) $z_{\bar{G}} = 1$. Therefore, the system (35)-(36) has a unique integer solution.

2) $r = 1$: By (14) and (29), $r = 1$ implies $z_{\bar{G}} = 0$. Then, the system (35)-(36) reduces to

$$z_{\bar{G}-1} + z_{\bar{G}-2} = 1, \tag{37}$$

$$(\bar{G}-1)z_{\bar{G}-1} + (\bar{G}-2)z_{\bar{G}-2} = \bar{G} - 1, \tag{38}$$

with a unique integer solution $z_{\bar{G}-1} = 1$, $z_{\bar{G}-2} = 0$.

3) $r = 2$: By (14) and (29), $r = 2$ implies $z_{\bar{G}} = 0$, and the resulting system of equations

$$z_{\bar{G}-1} + z_{\bar{G}-2} = 1, \tag{39}$$

$$(\bar{G}-1)z_{\bar{G}-1} + (\bar{G}-2)z_{\bar{G}-2} = \bar{G} - 2. \tag{40}$$

has a unique integer solution $z_{\bar{G}-1} = 0$ and $z_{\bar{G}-2} = 1$.

Therefore, in models P_B, P_E with equations (29) and (30) and conditions $i)-iii)$ satisfied for some $(i',t') \in \mathcal{I}\times\mathcal{T}$, the system (11) and (16) for (i',t') has a unique solution and this solution is integer even if the integrality condition on the $z_{i't'k}$ variables is relaxed for (i',t') and $\forall k \in \mathcal{K}(i',t')$. □

Acknowledgment

This research was funded by Rio Tinto and by the NSERC–Engage and MITACS-Accelerate programs.

References

[1] A. Froger, M. Gendreau, J.E. Mendoza, É. Pinson, L.M. Rousseau, "Maintenance Scheduling in the Electricity Industry: A Literature Review", *Eur. J. Oper. Res.*, vol. 251, no. 3, pp. 695-706, 2016.
[2] Y. Feng, P. Li, H. Wang, "Hydro-thermal generator maintenance scheduling accommodating both randomness and fuzziness", *4th Int. Conf. on Electric Utility Deregulation and Restructuring and Power Tech. IEEE*, pp. 734-741, 2011.
[3] W.K. Foong, A.R. Simpson, H.R. Maier, S. Stolp, "Ant colony optimization for power plant maintenance scheduling optimization—a five-station hydropower system", *Annals of Operations Research*, vol. 159, no. 1, pp. 433-450, 2008.
[4] I. Kuzle, H. Pandžić, M. Brezovec, "Hydro Generating Units Maintenance Scheduling Using Benders Decomposition", *Tehnički vjesnik*, vol. 17, no. 2, pp. 145-152, 2010.
[5] S. Perez-Canto, "Application of Benders' decomposition to power plant preventive maintenance scheduling", *Eur. J. Oper. Res.*, vol. 184, no. 2, pp. 759–777, 2008.
[6] E.C. Finardi and E.L. da Silva, "Solving the hydro unit commitment problem via dual decomposition and sequential quadratic programming", *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp.835-844, 2006.
[7] A. Arce, T. Ohishi, and S. Soares, "Optimal Dispatch of Generating Units of the Itaipú Hydroelectric Plant", *IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 154-158, 2002.
[8] J.P.S. Catalão, S.J.P.S. Mariano, and V.M.F. Mendes, and L.A.F.M. Ferreira, "Scheduling of head-sensitive cascaded hydro systems: A nonlinear approach", *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 337-346, 2009.
[9] A.J. Conejo, J.M. Arroyo, J. Contreras, F.A. Villamor, "Self-scheduling of a hydro producer in a pool-based electricity market", *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1265-1272, 2002.
[10] A. Borghetti, C. D'Ambrosio, A. Lodi, S. Martello, "An MILP approach for short-term hydro scheduling and unit commitment with head-dependent reservoir", *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1115-1124, 2008.
[11] A.L. Diniz, and M.E.P. Maceira, "A four-dimensional model of hydro generation for the short-term hydrothermal dispatch problem considering head and spillage effects", *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1298-1308, 2008.
[12] S. Séguin, P. Côté, C. Audet, "Self-scheduling short-term unit commitment and loading problem", *IEEE Trans. Power Syst.* vol. 31, no. 1, pp. 133-142, 2016.
[13] R. Bixby, E. Rothberg, "Progress in computational mixed integer programming—A look back from the other side of the tipping point", *Annals of Operations Research.* vol. 149, no. 1, pp. 37-41, 2007.
[14] G. Papazafeiropoulos, 2017-02-15, "Computational Geometry Toolbox", https://goo.gl/TYuv3E
[15] K. Fukuda, 2017-02-15, "Polyhedral Computation Software", https://www.inf.ethz.ch/personal/fukudak/soft/soft.html
[16] Fico, "Xpress-BCL Reference manual", Release 4.8.1, 2015.
[17] E.D. Dolan, and J.J. Moré, "Benchmarking optimization software with performance profiles", *Mathematical Programming*, vol. 91, no. 2, pp. 201-213, 2002.